Title

Functional interplay between NTP leaving group and base pair recognition during RNA polymerase II nucleotide incorporation revealed by methylene substitution.

Journal

Nucleic acids research, 44(8)

ISSN

0305-1048

Authors

Hwang, Candy S
Xu, Liang
Wang, Wei
et al.

Publication Date

2016-05-01

DOI

10.1093/nar/gkw220

Peer reviewed

3820–3828 *Nucleic Acids Research, 2016, Vol. 44, No. 8* *Published online 7 April 2016*
doi: 10.1093/nar/gkw220

Functional interplay between NTP leaving group and base pair recognition during RNA polymerase II nucleotide incorporation revealed by methylene substitution

Candy S. Hwang[1,†], Liang Xu[2,†], Wei Wang[2], Sébastien Ulrich[3,4], Lu Zhang[5], Jenny Chong[2], Ji Hyun Shin[2], Xuhui Huang[5], Eric T. Kool[3], Charles E. McKenna[1,*] and Dong Wang[2,*]

[1]Department of Chemistry, University of Southern California, Los Angeles, CA 90089-0744, USA, [2]Department of Cellular and Molecular Medicine, School of Medicine; Skaggs School of Pharmacy and Pharmaceutical Sciences, The University of California, San Diego, La Jolla, CA 92093-0625, USA, [3]Department of Chemistry, Stanford University, Stanford, CA 94305-5017, USA, [4]Institut des Biomolécules Max Mousseron (IBMM), UMR 5247 CNRS, Université Montpellier, ENSCM, Ecole Nationale Supérieure de Chimie de Montpellier, 8 Rue de l'Ecole Normale, 34296 Montpellier cedex 5, France and [5]Department of Chemistry, Division of Biomedical Engineering, Center of Systems Biology and Human Health, School of Science and Institute for Advance Study, The Hong Kong University of Science and Technology, Clear Water Bay, Kowloon, Hong Kong

Received October 22, 2015; Revised March 18, 2016; Accepted March 21, 2016

ABSTRACT

RNA polymerase II (pol II) utilizes a complex interaction network to select and incorporate correct nucleoside triphosphate (NTP) substrates with high efficiency and fidelity. Our previous 'synthetic nucleic acid substitution' strategy has been successfully applied in dissecting the function of nucleic acid moieties in pol II transcription. However, how the triphosphate moiety of substrate influences the rate of P-O bond cleavage and formation during nucleotide incorporation is still unclear. Here, by employing β,γ-bridging atom-'substituted' NTPs, we elucidate how the methylene substitution in the pyrophosphate leaving group affects cognate and non-cognate nucleotide incorporation. Intriguingly, the effect of the β,γ-methylene substitution on the non-cognate UTP/dT scaffold (∼3-fold decrease in k_{pol}) is significantly different from that of the cognate ATP/dT scaffold (∼130-fold decrease in k_{pol}). Removal of the wobble hydrogen bonds in U:dT recovers a strong response to methylene substitution of UTP. Our kinetic and modeling studies are consistent with a unique altered transition state for bond formation and cleavage for UTP/dT incorporation compared with ATP/dT incorporation. Collectively, our data reveals the functional interplay between NTP triphosphate moiety and base pair hydrogen bonding recognition during nucleotide incorporation.

INTRODUCTION

RNA polymerase II (pol II) is the central enzyme responsible for the synthesis of mRNA during eukaryotic gene transcription (1). It is essential for pol II to catalyze nucleotide addition with high efficiency and fidelity. Previous studies revealed that a key conserved motif, the trigger loop (TL), plays important roles in substrate selection, catalysis and translocation during each nucleotide addition cycle in the transcription elongation phase (2–9). Upon correct substrate binding, the TL undergoes important conformational changes, switching from inactive open conformations to an active closed conformation (2–4,6–7). This closure of the TL stabilizes the correctly bound substrate and facilitates subsequent nucleotide addition reaction (2–9). The pol II TL, bridge helix (BH), along with other active site residues, the nucleoside triphosphate (NTP) substrates and the RNA–DNA hybrid form a well-organized substrate recognition network to ensure high transcriptional fidelity (Figure 1A) (10–12). Multiple types of non-covalent interactions, such as base stacking, hydrogen-bonding, hydrophobic interaction and salt bridge interactions, are involved in substrate recognition (Figure 1A) (13). Dissecting individual contributions from these interactions within

*To whom correspondence should be addressed. Tel: +1 858 822 5561; Fax: +1 858 822 1953; Email: dongwang@ucsd.edu
Correspondence may also be addressed to Charles E. McKenna. Tel: +1 213 740 7007; Fax: +1 213 740 0930; Email: mckenna@usc.edu
† These authors contributed equally to the paper as first authors.

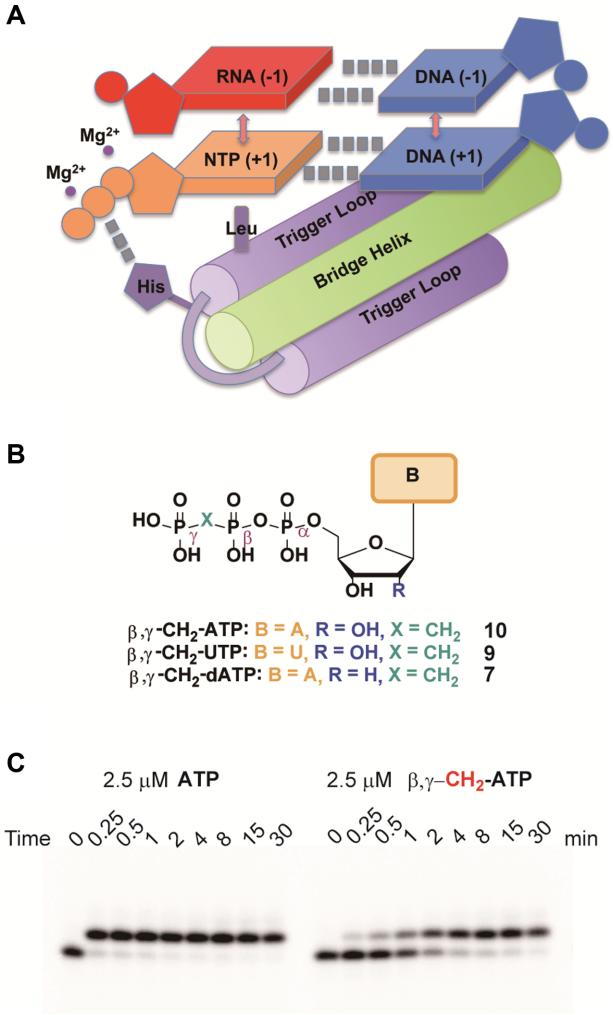

Figure 1. β,γ-CH_2-(d)NTPs can be recognized and incorporated by RNA pol II. (**A**) Interaction network in pol II catalytic site. (**B**) Chemical structures of β,γ-CH_2 (d)NTPs. (**C**) Gel analysis of nucleotide incorporation of *wt* ATP and its analog (2.5 μM). Time points were 15 s, 30 s, 1 min, 2 min, 4 min, 8 min, 15 min and 30 min.

this sophisticated network is pivotal to understanding the overall molecular mechanism of pol II transcription and the molecular basis of how DNA modifications and lesions affect transcription fidelity (14–17).

Synthetic nucleic acid analogs are powerful tools to probe interactions between nucleic acids and pol II by 'substituting' individual nucleic acid moieties or ablating specific interactions (18–20). These nucleic acid analogs enabled us to investigate the individual contributions of the nucleobase and sugar moieties to pol II transcription. These studies have advanced our understanding of how the intrinsic structural features of nucleic acid moieties are recognized and how specific interactions are involved in substrate selection and incorporation during pol II transcription (21–24). However, the interactions involved at the triphosphate moiety of NTPs during nucleotide selection and incorporation have not yet been extensively explored. Investigations targeting the triphosphate moiety can provide further mechanistic insights into how pol II catalyzes the chemical bond cleavage and formation during nucleotide addition.

Deoxynucleoside triphosphate (dNTP) analogs were applied to probe the mechanisms of several nucleic acid enzymes (25–27). The synthesis of these dNTP analogs involves replacing the bridging β,γ-oxygen with a methylene group (28), resulting in a non-hydrolysable analog for enzymes that cleave the terminal phosphate (Figure 1B) (29,30). For enzymes such as DNA pol β that release pyrophosphate ion (PP_i) after nucleotidyl transfer (25–27), these β,γ-CXY-dNTP analogs function as substrate analogs with leaving group properties that can be tuned by substituents on the methylene carbon (31–33). The bridging oxygen substituted by the less electronegative methylene group increases the leaving group's basicity relative to pyrophosphate and thereby provides a method to examine the rate-determining step (RDS) of DNA pol β (31–33). Here, by employing β,γ-CH_2 modified (d)NTPs (Figure 1B), we explore how RNA pol II recognizes modified triphosphate groups and how the chemical step affects nucleotide incorporation efficiency.

MATERIALS AND METHODS

Synthesis of β,γ-CH_2-(d)NTP nucleotide analogs

Tetraisopropyl methylenebis(phosphonate) **1** was a gift from Rhodia. Nucleoside 5′-monophosphates were purchased from Sigma-Aldrich (UMP, **4**) and Chem-Impex International (dAMP, **3**). Nucleoside 5'-monophosphate morpholidates ((d)NMP-morpholidate, (d)NMP 5′-M, **5**,**6**) were prepared according to published procedure (26,34–35). β,γ-CH_2-ATP (**10**) is commercially available from Sigma-Aldrich (Figure 1B). β,γ-CH_2-dATP and β,γ-CH_2-UTP (compounds **7** and **9**, respectively, Figure 1B) (36) were synthesized by the reaction of a tributylammonium salt of methylenebis(phosphonic acid) (**2**) with 5′-morpholidate (**5** or **6**) of the respective 5′-(d)NMP (26,35). All other reagents were purchased from commercial sources and used as obtained, unless specified otherwise. All compounds were further purified by dual-pass preparative High pressure liquid chromatography (HPLC) methods described previously (35,37) and characterized by UV, ^{31}P and ^{1}H NMR and MS (for spectra and HPLC data refer to Supplementary Figures S2–13 and Table S1 in the Supplementary Information). ^{1}H and ^{31}P NMR spectra were obtained on VNMRS-500 and VNMRS-600 3-Channel NMR spectrometers. The stability of these analogs was confirmed by LC-MS (Supplementary Table S2). Compounds **7, 9, 10** were used for bioassays with RNA pol II (Figure 1B).

In vitro transcription assays

RNA pol II was purified from *Saccharomyces cerevisiae* as previously described (2,38). The DNA template and non-template oligonucleotides were purchased from IDT. RNA primers were purchased from TriLink Biotechnologies and radiolabeled using [γ-^{32}P] adenosine triphosphate (ATP) and T4 Polynucleotide Kinase (NEB). The ultrapure NTPs were purchased from Affymetrix.

The pol II elongation complexes for transcription assays were assembled using established methods (21,22). Briefly,

an aliquot of 5′-^{32}P-labeled RNA (10 μM) was annealed with a 1.5-fold amount of template DNA (15 μM) and 2-fold amount of non-template DNA (20 μM) to form the RNA/DNA scaffold (final stock concentration: 1 μM, defined by RNA concentration) in elongation buffer (20 mM Tris–HCl (pH 7.5), 40 mM KCl, 5 mM $MgCl_2$). An aliquot of the annealed scaffold of RNA/DNA (50 nM) was then incubated with a 4-fold amount of pol II (200 nM) at room temperature for 10 min to ensure the formation of a pol II elongation complex. The pol II elongation complex is ready for *in vitro* transcription upon mixing with equal volumes of NTP solution of various concentrations. Final reaction concentrations after mixing were 25 nM scaffold, 100 nM pol II, 5 mM DTT, 5 mM $MgCl_2$, 40 mM KCl, 20 mM Tris–HCl (pH = 7.5) and NTP. The quenched products were analyzed by 16% denaturing urea polyacrylamide gel electrophoresis in TBE (Tris/Borate/EDTA) buffer and visualized using a storage phosphor screen and Pharos FX imager (Bio-Rad). All *in vitro* transcription assays described in this manuscript were repeated two to four times.

Single turnover nucleotide incorporation assays

The assay was carried out as previously described (21,22). Briefly, nucleotide incorporation assays were conducted by pre-incubating 50 nM scaffold with 200 nM pol II for 10 min in elongation buffer at 22°C. The pre-incubated enzyme:scaffold complex was then mixed with an equal volume of solution containing 40 mM KCl, 20 mM Tris–HCl (pH 7.5), 10 mM DTT, 10 mM $MgCl_2$ and 2-fold concentrations of various nucleotides. Final reaction concentrations after mixing were 25 nM scaffold, 100 nM pol II, 5 mM $MgCl_2$ and various nucleotide concentrations in elongation buffer. Reactions were quenched at various times by the addition of one volume of 0.5 M ethylenediaminetetraacetic acid (pH 8.0), and analyzed as described above.

Kinetic data analysis

Non-linear-regression data fitting was performed using Prism 6. The time dependence of product formation was fit to a one-phase association Equation (1) to determine the observed rate (k_{obs}). The substrate concentration dependence was fit to a hyperbolic Equation (2) to obtain values for the maximum rate of NTP incorporation (k_{pol}) and apparent K_d ($K_{d,app}$) governing NTP binding.

$$[\text{Product}] = Ae^{(-k\text{obs t})} + C \quad (1)$$

$$k_{obs} = k_{pol}[\text{Substrate}]/(K_{d,app} + [\text{Substrate}]) \quad (2)$$

Representative data and kinetic fittings are shown in Supplementary Figures S14 and 15. The specificity constant was determined by $k_{pol}/K_{d,app}$. Discrimination was calculated as the ratio of specificity constants governing incorporation of cognate substrate over non-cognate substrate as described (21,22).

Molecular modeling and energy minimization

The pol II elongation complex model was built based on the crystal structures of the pol II complex (39) with bound cognate ATP (PDB ID: 4BY1). The model of pol II elongation complex with a non-cognate UTP opposite dT was obtained in the following procedure: first, the crystal coordinates of two stable U:U pairs were obtained from the crystal structure of an RNA duplex (PDB ID: 4U38) (40). These U:U pairs form specific wobble hydrogen bonding patterns. One of the UMPs was then substituted by UTP (from the PDB ID: 2NVZ) to generate a UTP:U pair (2). The nucleobase of the UMP (corresponding to the template strand) in UTP:U pairs was superimposed with the dT at the $i + 1$ site of template strand in pol II complex using COOT (41). This superposition allowed us to define the potential positions of UTP in the pair of UTP:dT at pol II active site. All structure figures were rendered in PyMOL (http://www.pymol.org).

Energy minimization of described model above was then performed. Protein subunits Rpb4 (chain D) and Rpb7 (chain G) were removed. Missing residues were built using Modeller (42,43). The missing side chains of amino acids were fixed by the 'what if' suite (44). Protonation states of the titratable residues were predicted using Propka (45,46) in Pdb2pqr package (47). The Amber99sb force field parameters (48) were chosen to simulate protein, nucleic acids and ions. Amber99sb force field parameters for uridine were employed for the base and sugar ring of UTP. Amber-compatible polyphosphate parameters were developed from Meagher *et al.* (49) and were used for the triphosphate moiety of UTP.

Each system was solvated in a triclinic box with the box edges at least 10 Å away from the Pol II surface. Then 85 sodium ions were added to make the system electrically neutral. The whole simulation model contained ∼450 000 atoms, including ∼129 000 TIP3P water molecules (50). Afterward, the package Gromacs 5.0 (51,52) was used for two-phase energy minimizations: a 5000-step energy minimization with the steepest descent algorithm was performed first with positional restraint (i.e. a force constant of 10 kJ mol^{-1} $Å^{-2}$) on the heavy atoms of the nucleotides and UTP; followed by another 10 000-step energy minimization performed for the whole system. The long-range electrostatic interactions beyond the cut-off of 12 Å were treated with the Particle-Mesh Ewald method (53). The van der Waals interactions were smoothly switched off between 10 and 12 Å.

RESULTS

Cognate nucleotide incorporation

During transcription, pol II catalyzes bond cleavage at the P_α-O-P_β site of an NTP substrate and bond formation of NMP with the 3′-OH of RNA terminus. As a result, the RNA transcript is extended by 1 nt at its 3′-end and a PP_i, the leaving group, is subsequently released from the pol II active site. Substitution of the β,γ-bridging oxygen in the triphosphate with methylene (Figure 1B) (28) leads to a poorer leaving group than PP_i. Such nucleotide analogs allow us to determine the incorporation rates of the natural (or 'wild-type' (*wt*)) NTP and β,γ-CH_2 modified NTPs (β,γ-CH_2-ATP (10), β,γ-CH_2-dATP (7), and β,γ-CH_2-UTP (9)) and to probe the mechanisms of pol II catalysis and transcriptional fidelity control.

In vitro transcription assays were first performed to evaluate substrate incorporation efficiencies for *wt* ATP and β,γ-CH_2-ATP (Figure 1B and C). As shown in Figure 1C, β,γ-CH_2-ATP can be recognized by pol II as a substrate and becomes incorporated into the RNA primer via a Watson–Crick (W–C) base pair with dT. However, β,γ-CH_2 substitution in ATP greatly reduces the nucleotide incorporation rate by pol II in comparison to *wt* ATP (Figures 1C and 2B).

To quantitatively identify the effect of the bridging β,γ-CH_2 substitution, single-turnover pre-steady-state kinetic assays were performed to measure the kinetic parameters of pol II nucleotide incorporation, specifically k_{pol} and $K_{d,app}$, and the ratio of $k_{pol}/K_{d,app}$ to calculate the substrate specificity constant of pol II (Supplementary Table S3). The kinetic analysis revealed that the incorporation of β,γ-CH_2-ATP was significantly slower than *wt* ATP, with a ∼130-fold decrease in k_{pol}. Interestingly, the binding affinity ($K_{d,app}$) of β,γ-CH_2-ATP in the dT template did not vary substantially from *wt* ATP, with both values in the moderately low micromolar range. As a result, nucleotide incorporation specificity $k_{pol}/K_{d,app}$ was greatly reduced (∼50-fold). Thus, in the cognate ATP/dT scaffold, the CH_2 substitution of the bridging oxygen atom results in decreased catalytic efficiency, indicating that the CH_2 substitution substantially slows down the 'chemistry step' (P-O bond cleavage of the triphosphate moiety between P_α-O-P_β and formation of NMP with the 3′-OH of RNA terminus) and the 'chemistry step' becomes the RDS of the overall mechanism (31,32).

Non-cognate nucleotide incorporation

Since the cognate (matched) nucleotide incorporation is negatively affected by the β,γ-CH_2 substitution, it is of interest to determine whether this substitution also alters non-cognate nucleotide incorporation, where the nucleobase or sugar moiety of the incoming NTP is incorrect (27). We first investigated the CH_2 substitution effect on non-cognate (mismatched) substrate incorporation using UTP as an example (Figure 1B). Intriguingly, β,γ-CH_2-UTP exhibits only a 3-fold decrease in k_{pol} and a 9-fold decrease in specificity constant ($k_{pol}/K_{d,app}$) in comparison with *wt* UTP (Figure 2 and Supplementary Table S3). This pattern is strikingly distinct from that of the cognate ATP/dT nucleotide incorporation.

We then investigated the CH_2 substitution effect on non-cognate nucleotide dATP incorporation. Interestingly, β,γ-CH_2-dATP exhibits ∼100-fold decrease in k_{pol} and ∼20-fold in specificity constant ($k_{pol}/K_{d,app}$) in comparison with *wt* dATP (Figure 2 and Supplementary Table S3). This pattern is comparable with the results of β,γ-CH_2-ATP for the cognate ATP/dT nucleotide incorporation.

These results indicate that both *wt* and β,γ-CH_2 modified nucleotide analogs can be recognized by pol II as substrates. However, replacement of pyrophosphate leaving group with a methylenebisphosphonate group has a profound effect on nucleotide addition efficiency. Specifically, the methylene modification effects on nucleotide incorporation are sensitive to the nature of substrate base-pairing interactions but not sugar identity.

Trigger loop-dependent substrate recognition and incorporation

The TL plays key roles in controlling efficient pol II nucleotide selection and catalysis (2–3,6–7). The TL undergoes a conformational change from an inactive open state to a closed active state in the presence of correct W–C base pairs with the DNA template strand, allowing efficient and accurate nucleotide addition (6,7). To further understand the effect of methylene substitution of (d)NTPs on pol II nucleotide-binding induced TL conformational changes and the extent of TL-dependent nucleotide incorporation, we measured the α-amanitin effect index (AEI). AEI is the ratio of nucleotide incorporation efficiencies in the absence and presence of α-amanitin. This was determined for ATP/dT, dATP/dT and UTP/dT and their CH_2 modified analogs to assess the involvement of active TL during substrate incorporation (Supplementary Table S4) (23). α-amanitin is a potent pol II-specific inhibitor that binds nearby BH and TL, and traps TL in an open inactive conformation, preventing it from closing to an active closed conformation (3,54). As a result, AEI-sensitivity in NTP incorporation indicates the extent of TL involvement (requires the closed active TL conformation) during nucleotide incorporation. The TL is actively involved in substrate recognition for cognate nucleotide incorporation and reaches fully closed state upon cognate substrate binding, and hence, it results in high AEI values (∼80–100). In contrast, the TL fails to reach its fully closed state effectively in the non-cognate nucleobase or sugar scenario, and thus smaller AEI values are observed (∼1 or ∼10, respectively) (23).

Indeed, as shown in Figure 3 and Supplementary Table S4, the *wt* ATP revealed a high AEI value (∼80) for the cognate AMP incorporation, indicating that the TL recognizes the *wt* substrate. It is important to note that a high AEI value was also observed with β,γ-CH_2-ATP (∼80). These results establish that β,γ-methylene substitution in ATP does not interfere with TL closure during matched AMP incorporation. In the non-cognate nucleotide incorporation (UTP/dT) scenario, similar AEI values (<5) were observed for both UTP and β,γ-CH_2-UTP incorporation, indicating that the TL remains in an inactive conformation for both *wt* UTP and β,γ-CH_2-UTP analog incorporation (Figure 3). Similarly, we observed comparable AEI values (∼5–10) for both *wt* dATP and β,γ-CH_2-dATP for dATP/dT incorporation (non-cognate sugar). Taken together, methylene substitution of the β,γ-bridge atom of the triphosphate moiety does not modify TL open-close conformational profile and its role in the selection of cognate substrate (ATP) over non-cognate substrate (UTP and dATP).

Effect of wobble base-pairing on UMP incorporation

The observation that TL recognizes β,γ-CH_2-NTPs and NTPs in a similar manner indicates that they are appropriate probes to investigate the mechanism of pol II catalysis by modulating the leaving group and the chemistry step. One intriguing finding is that the methylene substitution generates a significant difference in k_{pol} for cognate AMP incorporation but a minor difference for UMP incorporation.

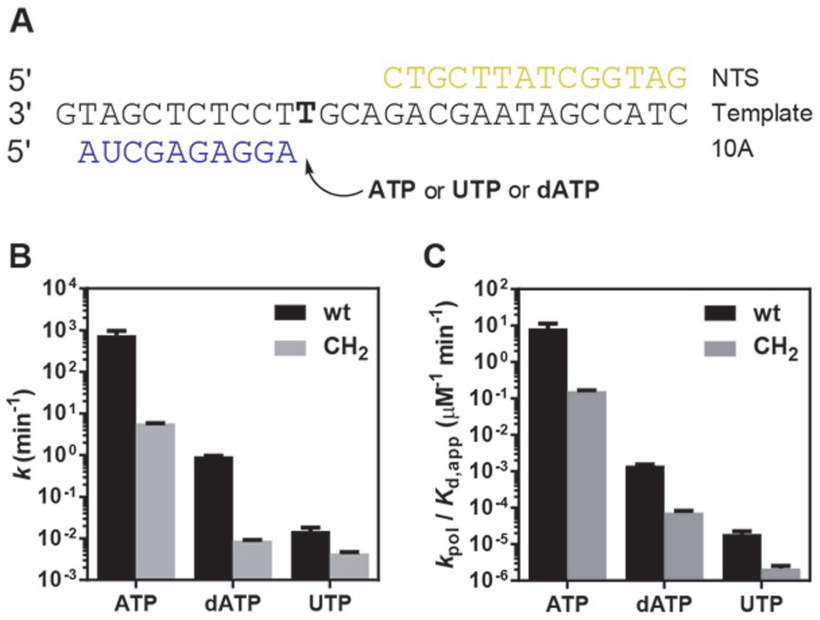

Figure 2. Kinetic effects of varying the β,γ-bridging atom in NTPs during RNA pol II incorporation. (**A**) Scaffold used for studying incorporation efficiency of cognate nucleotide (AMP/dT) and non-cognate nucleotides (dAMP/dT and UMP/dT). Catalytic rates (k_{pol}) (**B**) and specificity constants (k_{pol}/K_d) (**C**) of *wt* and β,γ-CH_2-(d)NTPs.

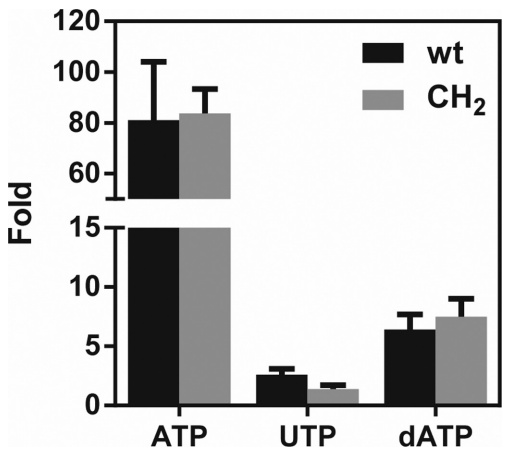

Figure 3. α-Amanitin effect index reveals that β,γ-CH_2 substitution doesn't interfere with patterns of TL dependence for nucleotide incorporation.

Since UTP forms a wobble-base pair with dT that potentially positions its triphosphate moiety in a different binding environment from that of ATP, we were interested in investigating whether these distinct methylene substitution effects rely on the patterns of hydrogen bonding to the incipient pair. To this end, we introduced the hydrogen-bonding deficient mutant (dF), a non-polar isostere of dT (55), to ablate hydrogen-bonding interactions between base pairs (21,23) and measured the β,γ-CH_2 substitution effects for both AMP and UMP incorporation on the dF template (Figure 4 and Supplementary Table S3).

For the incipient ATP/dF pair, methylene substitution also resulted in ∼70-fold decrease in k_{pol}, which is comparable to the result observed for the ATP/dT system. This indicates that the elimination of W–C base pair hydrogen bonding in the ATP/dF scaffold did not significantly alter the strong methylene substitution effect relative to the ATP/dT scaffold. It is also very interesting to note that we observe the similar methylene substitution effects from dT and dF scaffold, even though the initial *wt* ATP incorporation rate for dF scaffold is substantially (∼10^4) lower than that for dT scaffold, indicating that the effect of methylene substitution does not rely on the initial rate values. Interestingly, unlike the UTP/dT pair, we observed that methylene substitution greatly reduced incorporation rate in the UTP/dF setup. Notably, this stronger methylene substitution effect in the UTP/dF pair was due to the substantial increase in the incorporation rate for the natural UTP/dF case, while the incorporation rate for methylene substituted UTP stayed in a similar level in these two pairs.

The methylene substitution effects on NTP incorporation were examined in a systematic manner and summarized in Figure 5 and Supplementary Table S3. Methylene substitution greatly reduced incorporation of ATP for both the dT and dF templates, indicating that the strong methylene substitution effect does not rely on W–C hydrogen bonding between ATP and the matched DNA template. In sharp contrast, methylene substitution of non-cognate UTP incorporation leads to ∼3-fold decrease for UTP/dT system, whereas it causes ∼80-fold reduction for UTP/dF incorporation. We found that the methylene substitution effect on the UTP/dT pair is distinct from the rest of the groups (Figure 5 and Supplementary Table S3). These intriguing findings indicate that the removal of wobble pair hydrogen bonding between UTP/dT pairing abolishes its unique response to methylene substitution in the triphosphate moiety.

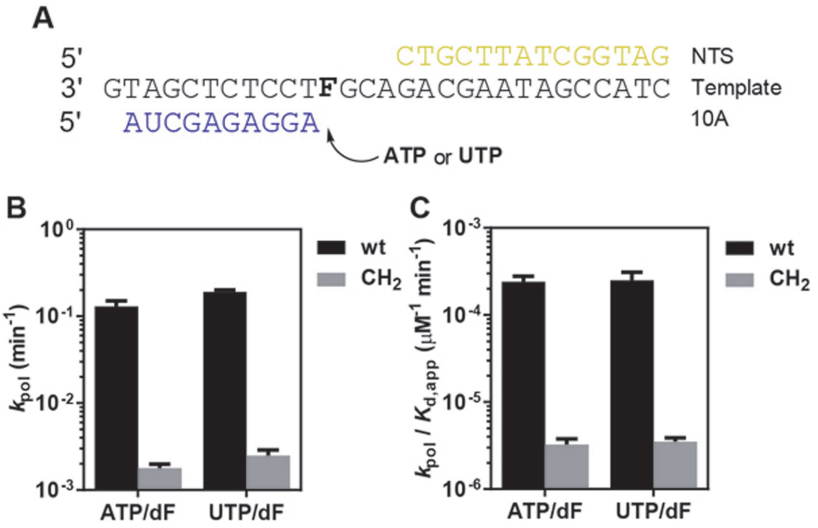

Figure 4. Kinetic effects of alteration of β,γ-bridging atoms in NTPs during RNA pol II incorporation in hydrogen bond deficient scaffolds. (**A**) Scaffold used for studying incorporation efficiency in the hydrogen-bonding deficient template (dF). Catalytic rates (k_{pol}) (**B**) and specificity constants (k_{pol}/K_d) (**C**) of *wt* NTPs, and β,γ-CH_2-NTPs with the dF template.

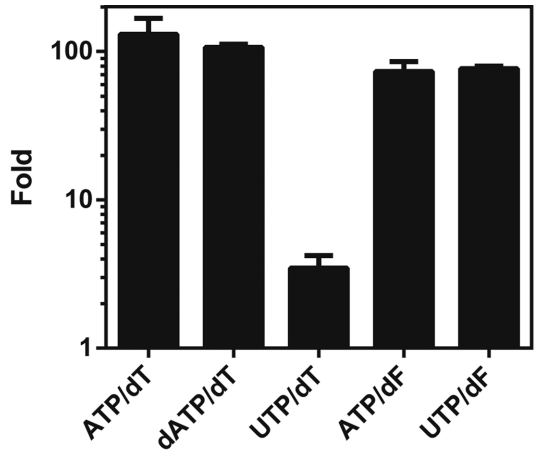

Figure 5. CH_2 substitution effect on k_{pol} in different scaffolds.

To gain further structural insight into how UTP binds in the pol II active site, we performed molecular modeling based on a previously published high resolution wobble base-pairing crystal structure (40). Two possible wobble-base pairing patterns were obtained for UTP:dT in the context of the RNA:DNA hybrid duplex (Supplementary Figure S16) (40). When we superimposed these two states in the pol II active site (2), we found that only one UTP:dT wobble-pair pairing was energetically preferred (Figure 6A), whereas the other wobble-pair causes strong steric clash with neighboring pol II residues (Figure 6B). Further energy minimization of the favorable wobble-base pairing form in Figure 6A reveals that UTP is well tolerated at pol II catalytic site (Supplementary Figure S17A). Interestingly the uracil base of incoming UTP may likely form additional electrostatic interactions with the main chain of nearby pol II residues in fork loop 3, which we previously reported to play a recognition role toward modified cytosine in the template (56). In parallel, we also applied the same energy minimization procedure for the 'sterically clashed' wobble-base pairing form in Figure 6B. Although the strong steric clash was partially resolved by pushing both the protein residue and the uracil base of UTP away from each other, it is not energetically favorable for closed contact between the hydrophobic residue (P448) and the hydrophilic uracil base (Supplementary Figure S17B), Hence, we propose that wobble base paring of UTP:dT in Figure 6A (and Supplementary Figure S17A) may represent the major form during UTP/dT incorporation. In this energy-favored conformation, the UTP:dT wobble geometry causes UTP to shift away from the canonical position (nucleotide addition site). Such a positional shift also causes the triphosphate moiety of UTP to move away from its canonical binding site, which contains several positively charged interacting residues. Thus, this positional shift of the triphosphate may prevent nucleotidyl transfer reaction, disfavor the stabilization of the transition state (2,57) and ultimately results in the unique CH_2 substitution pattern observed for the UTP/dT pair.

DISCUSSION

RNA pol II possesses high nucleotide incorporation efficiency and fidelity during transcriptional elongation to ensure accurate and timely genetic information transfer. Application of a 'synthetic nucleic acid substitution' strategy has revealed several unique contributions from the nucleic acid nucleobase and sugar backbone moieties (21–23). In this study, we utilized β,γ-CH_2-NTP 'analogs' as probes to dissect the relationship between 'triphosphate' moiety and the mechanism of nucleotide incorporation. These probes provide unique information pertaining to the roles of base-pairing and how the leaving group affects cognate (matched) and non-cognate nucleotide addition by pol II.

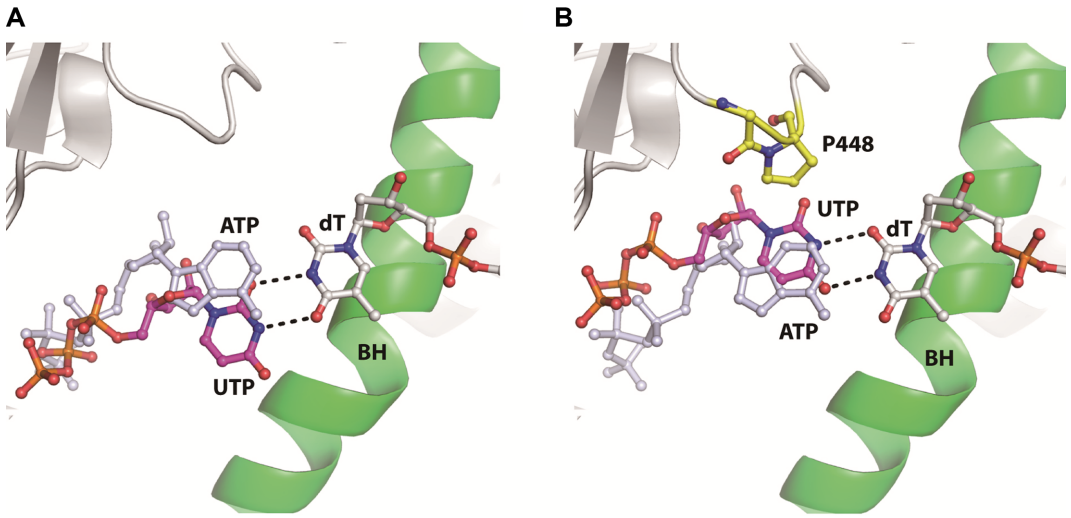

Figure 6. Molecular modeling of UTP binding opposite dT template at pol II active site. (**A**) UTP forms a wobble bonding pair with dT template, resulting in the misalignment of its triphosphate group compared to cognate ATP recognition. (**B**) An alternative form of UTP:dT pairing In this form, the UTP shifts toward the minor groove of the RNA/DNA hybrid, leading to steric clash between pol II residues and the incoming nucleotide. The two potential base pairing behaviors were modeled based on the previously reported U:U pair (42).

It was reported that the chemistry step is not the rate-limiting step for pol II nucleotide incorporation. Rather, non-chemistry steps, such as TL conformational changes upon substrate binding and translocation steps, are likely to be rate-limiting steps (58–61). On the other hand, direct evidence for dissecting and comparing the relative time-scales of the chemistry step with those non-chemistry rate-limiting steps is lacking. Here we reported a group of β,γ-CH_2-NTP analogs that act as 'functional mutants' to specifically modulate the chemistry step without disrupting the key recognition process of TL closure.

To our surprise, we found that CH_2 substitution resulted in ∼130-fold decrease in k_{pol} in the cognate ATP/dT scaffold. The effect on the rate of chemical bond cleavage and formation is significantly higher than previously observed for DNA pol β (31,32). Indeed, comparison of methylene substitution on RNA pol II and DNA polymerase catalyzed DNA-dependent substrate incorporation reveals both similarities and differences. In terms of similarities, we find that methylene substitution causes a decrease in incorporation efficiency in both enzymes. The weaker leaving group resulting from CH_2 substitution is expected to slow down P-O bond cleavage of the triphosphate and substrate incorporation. These results in a shift of the overall rate-limiting step to the 'chemistry' step. However, RNA pol II is more sensitive to methylene substitution with around ∼100-fold rate decrease in cognate nucleotide incorporation, whereas DNA polymerase β is only affected by ∼10-fold. This result may suggest that RNA pol II active site is highly sensitive to electrostatic effect, in which variations in the basicity of pyrophosphate leaving group can greatly affect the interaction network, leading to a more significant response in the chemistry step than that of DNA polymerase β during nucleotide addition.

Another striking result is that the effect of the β,γ-CH_2 substitution on the non-cognate UTP/dT scaffold is significantly different from that of the matched ATP/dT scaffold. We found that methylene substitution leads to ∼130-fold less incorporation efficiency for the ATP/dT scaffold, but only causes ∼3-fold decrease in incorporation efficiency for the UTP/dT scaffold. This striking distinction likely unveils a unique altered transition state for bond formation and cleavage for UTP/dT incorporation compared with matched ATP/dT incorporation. Our structural modeling suggests that UTP likely adopts a distinct position (via wobble pair with dT template) in the active site with a different set of interacting pol II residues. Indeed, ablation of the wobble hydrogen bonds in the UTP/dF scaffold leads to a strong CH_2 substitution effect that resembles the matched ATP/dT pair (Figure 5).

Collectively, our data from comparative studies of *wt* NTPs and β,γ-CH_2-bridging NTP analogs reveals that pol II catalytic site is highly sensitive to variation in the basicity of the leaving group. RNA pol II transcriptional efficiency is not only determined by effective substrate recognition but also influenced by leaving group properties. Notably, we found two distinct states probed by these NTP analogs. One is the base cognate state, in which the interaction network is built on the canonical W–C base pair and the complementary shape for high incorporation efficiency. The other one is the mismatched UTP/dT form, in which a different interaction network is constructed induced by wobble UTP:dT hydrogen bonds. The uniqueness of UTP/dT incorporation may reveal how pol II takes advantage of UTP:dT wobble pairing to discriminate against UTP misincorporation. Future studies will be taken to obtain the direct structural evidence for this alternate UTP:dT binding and bond formation state and systematically investigate additional non-cognate incorporations using β,γ-CXY NTP analogs.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENT

The authors would like to thank Dr Boris A. Kashemirov (University of Southern California) for advice on the purification and analysis of compounds **7**, **9**, **10**.

FUNDING

Dana and David Dornsife College of the University of Southern California; National Institutes of Health [U19CA177547 to C.E.M, R01GM102362 to D.W., R01GM072705, R01GM106067 to E.T.K.]; National Science Foundation Graduate Research Fellowship [DGE-0937362 to C.S.H., in part]. Hong Kong Research Grant Council [HKUST C6009-15G to X.H.]. Funding for open access charge: NIH.

Conflict of interest statement. None declared.

REFERENCES

1. Svetlov,V. and Nudler,E. (2013) Basic mechanism of transcription by RNA polymerase II. *Biochim. Biophys. Acta*, **1829**, 20–28.
2. Wang,D., Bushnell,D.A., Westover,K.D., Kaplan,C.D. and Kornberg,R.D. (2006) Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell*, **127**, 941–954.
3. Kaplan,C.D., Larsson,K.-M. and Kornberg,R.D. (2008) The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. *Mol. Cell*, **30**, 547–556.
4. Kireeva,M.L., Nedialkov,Y.A., Cremona,G.H., Purtov,Y.A., Lubkowska,L., Malagon,F., Burton,Z.F., Strathern,J.N. and Kashlev,M. (2008) Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. *Mol. Cell*, **30**, 557–566.
5. Toulokhonov,I., Zhang,J., Palangat,M. and Landick,R. (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol. Cell*, **27**, 406–419.
6. Larson,M.H., Zhou,J., Kaplan,C.D., Palangat,M., Kornberg,R.D., Landick,R. and Block,S.M. (2012) Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of RNA polymerase II. *Proc. Natl. Acad. Sci. U.S.A.*, **109**, 6555–6560.
7. Huang,X., Wang,D., Weiss,D.R., Bushnell,D.A., Kornberg,R.D. and Levitt,M. (2010) RNA polymerase II trigger loop residues stabilize and position the incoming nucleotide triphosphate in transcription. *Proc. Natl. Acad. Sci. U.S.A.*, **107**, 15745–15750.
8. Feig,M. and Burton,Z.F. (2010) RNA polymerase II with open and closed trigger loops: active site dynamics and nucleic acid translocation. *Biophys. J.*, **99**, 2577–2586.
9. Zhang,J., Palangat,M. and Landick,R. (2010) Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat. Struct. Mol. Biol.*, **17**, 99–104.
10. Liu,X., Bushnell,D.A. and Kornberg,R.D. (2013) RNA polymerase II transcription: structure and mechanism. *Biochim. Biophys. Acta*, **1829**, 2–8.
11. Martinez-Rucobo,F.W. and Cramer,P. (2013) Structural basis of transcription elongation. *Biochim. Biophys. Acta*, **1829**, 9–19.
12. Lehmann,E., Brueckner,F. and Cramer,P. (2007) Molecular basis of RNA-dependent RNA polymerase II activity. *Nature*, **450**, 445–449.
13. Yuzenkova,Y., Bochkareva,A., Tadigotla,V.R., Roghanian,M., Zorov,S., Severinov,K. and Zenkin,N. (2010) Stepwise mechanism for transcription fidelity. *BMC Biol.*, **8**, 54–68.
14. Zhang,S., and Wang,D., (2013) Understanding the molecular basis of RNA polymerase II transcription. *Isr. J. Chem.*, **53**, 442–449.
15. Xu,L., Da,L., Plouffe,S.W., Chong,J., Kool,E. and Wang,D. (2014) Molecular basis of transcriptional fidelity and DNA lesion-induced transcriptional mutagenesis. *DNA Repair*, **19**, 71–83.
16. Walmacq,C., Cheung,A.C., Kireeva,M.L., Lubkowska,L., Ye,C., Gotte,D., Strathern,J.N., Carell,T., Cramer,P. and Kashlev,M. (2012) Mechanism of translesion transcription by RNA polymerase II and its role in cellular resistance to DNA damage. *Mol. Cell*, **46**, 18–29.
17. Xu,L., Wang,W., Chong,J., Shin,J.H., Xu,J. and Wang,D. (2015) RNA polymerase II transcriptional fidelity control and its functional interplay with DNA modifications. *Crit. Rev. Biochem. Mol. Biol.*, 1–17.
18. Chaput,J.C., Yu,H. and Zhang,S. (2012) The emerging world of synthetic genetics. *Chem. Biol.*, **19**, 1360–1371.
19. Benner,S.A. (2004) Understanding nucleic acids using synthetic chemistry. *Acc. Chem. Res.*, **37**, 784–797.
20. Pinheiro,V.B. and Holliger,P. (2012) The XNA world: progress towards replication and evolution of synthetic genetic polymers. *Curr. Opin. Chem. Biol.*, **16**, 245–252.
21. Kellinger,M.W., Ulrich,S., Chong,J., Kool,E.T. and Wang,D. (2012) Dissecting chemical interactions governing RNA polymerase II transcriptional fidelity. *J. Am. Chem. Soc.*, **134**, 8231–8240.
22. Xu,L., Plouffe,S.W., Chong,J., Wengel,J. and Wang,D. (2013) A chemical perspective on transcriptional fidelity: dominant contributions of sugar integrity revealed by unlocked nucleic acids. *Angew. Chem. Int. Ed. Engl.*, **52**, 12341–12345.
23. Xu,L., Butler,K.V., Chong,J., Wengel,J., Kool,E.T. and Wang,D. (2014) Dissecting the chemical interactions and substrate structural signatures governing RNA polymerase II trigger loop closure by synthetic nucleic acid analogues. *Nucleic Acids Res.*, **42**, 5863–5870.
24. Xu,L., Zhang,L., Chong,J., Xu,J., Huang,X. and Wang,D. (2014) Strand-specific (asymmetric) contribution of phosphodiester linkages on RNA polymerase II transcriptional efficiency and fidelity. *Proc. Natl. Acad. Sci. U.S.A.*, **111**, E3269–E3276.
25. McKenna,C.E., Kashemirov,B.A., Upton,T.G., Batra,V.K., Goodman,M.F., Pedersen,L.C., Beard,W.A. and Wilson,S.H. (2007) (R)-β,γ-Fluoromethylene-dGTP-DNA Ternary Complex with DNA Polymerase β. *J. Am. Chem. Soc.*, **129**, 15412–15413.
26. Wu,Y., Zakharova,V.M., Kashemirov,B.A., Goodman,M.F., Batra,V.K., Wilson,S.H. and McKenna,C.E. (2012) β,γ-CHF- and β,γ-CHCl-dGTP diastereomers: synthesis, discrete ^{31}P NMR signatures, and absolute configurations of new stereochemical probes for DNA polymerases. *J. Am. Chem. Soc.*, **134**, 8734–8737.
27. Oertell,K., Wu,Y., Zakharova,V.M., Kashemirov,B.A., Shock,D.D., Beard,W.A., Wilson,S.H., McKenna,C.E. and Goodman,M.F. (2012) Effect of β,γ-CHF- and β,γ-CHCl-dGTP Halogen Atom Stereochemistry on the transition state of DNA polymerase β. *Biochemistry*, **51**, 8491–8501.
28. Myers,T.C., Nakamura,K. and Flesher,J.W. (1963) Phosphonic acid analogs of nucleoside phosphate. I. The synthesis of 5'-adenylyl methylenediphosphonate, a phophonic acid analog of ATP. *J. Am. Chem. Soc.*, **85**, 3292–3295.
29. Boerner,R.J., Barker,S.C. and Knight,W.B. (1995) Kinetic mechanisms of the forward and reverse pp60c-srcT yrosine kinase reactions. *Biochemistry*, **34**, 16419–16423.
30. Elphick,L.M., Lee,S.E., Gouverneur,V. and Mann,D.J. (2007) Using chemical genetics and ATP analogues to dissect protein kinase function. *ACS Chem. Biol.*, **2**, 299–314.
31. Sucato,C.A., Upton,T.G., Kashemirov,B.A., Batra,V.K., Martinek,V., Xiang,Y., Beard,W.A., Pedersen,L.C., Wilson,S.H., McKenna,C.E. *et al.* (2007) Modifying the beta,gamma leaving-group bridging oxygen alters nucleotide incorporation efficiency, fidelity, and the catalytic mechanism of DNA polymerase beta. *Biochemistry*, **46**, 461–471.
32. Sucato,C.A., Upton,T.G., Kashemirov,B.A., Osuna,J., Oertell,K., Beard,W.A., Wilson,S.H., Florian,J., Warshel,A., McKenna,C.E. *et al.* (2008) DNA polymerase beta fidelity: halomethylene-modified leaving groups in pre-steady-state kinetic analysis reveal differences at the chemical transition state. *Biochemistry*, **47**, 870–879.
33. Oertell,K., Chamberlain,B.T., Wu,Y., Ferri,E., Kashemirov,B.A., Beard,W.A., Wilson,S.H., McKenna,C.E. and Goodman,M.F. (2014) Transition State in DNA polymerase β catalysis: rate-limiting chemistry altered by base-pair configuration. *Biochemistry*, **53**, 1842–1848.
34. Roseman,S., Distler,J.J., Moffatt,J.G. and Khorana,H.G. (1961) Nucleoside polyphosphates. XI.1 An improved general method for the synthesis of nucleotide coenzymes. Syntheses of uridine-5', cytidine-5' and guanosine-5' diphosphate derivatives. *J. Am. Chem. Soc.*, **83**, 659–663.
35. Hwang,C.S., Kashemirov,B.A. and McKenna,C.E. (2014) On the observation of discrete fluorine NMR Spectra for uridine

5′-β,γ-fluoromethylenetriphosphate diastereomers at basic pH. *J. Org. Chem.*, **79**, 5315–5319.
36. Mohamady,S. and Jakeman,D.L. (2005) An improved method for the synthesis of nucleoside triphosphate analogues. *J. Org. Chem.*, **70**, 10588–10591.
37. Upton,T.G., Kashemirov,B.A., McKenna,C.E., Goodman,M.F., Prakash,G.K.S., Kultyshev,R., Batra,V.K., Shock,D.D., Pedersen,L.C., Beard,W.A. *et al.* (2009) α,β-difluoromethylene deoxynucleoside 5′-triphosphates: a convenient synthesis of useful probes for DNA polymerase β structure and function. *Org. Lett.*, **11**, 1883–1886.
38. Wang,D., Bushnell,D.A., Huang,X., Westover,K.D., Levitt,M. and Kornberg,R.D. (2009) Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science*, **324**, 1203–1206.
39. Kinkelin,K., Wozniak,G.G., Rothbart,S.B., Lidschreiber,M., Strahl,B.D. and Cramer,P. (2013) Structures of RNA polymerase II complexes with Bye1, a chromatin-binding PHF3/DIDO homologue. *Proc. Natl. Acad. Sci. U.S.A.*, **110**, 15277–15282.
40. Sheng,J., Larsen,A., Heuberger,B.D., Blain,J.C. and Szostak,J.W. (2014) Crystal structure studies of RNA duplexes containing s2U:A and s2U:U base pairs. *J. Am. Chem. Soc.*, **136**, 13916–13924.
41. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D*, **60**, 2126–2132.
42. Eswar,N., Webb,B., Marti-Renom,M.A., Madhusudhan,M.S., Eramian,D., Shen,M.Y., Pieper,U. and Sali,A. (2006) Comparative protein structure modeling using Modeller. *Curr. Protoc. Bioinformatics*. John Wiley & Son, Inc, Chapter 5, Unit 5.6.
43. Fiser,A., Do,R.K.G. and Sali,A. (2000) Modeling of loops in protein structures. *Protein Sci.*, **9**, 1753–1773.
44. Chinea,G., Padron,G., Hooft,R.W.W., Sander,C. and Vriend,G. (1995) The use of position-specific rotamers in model-building by homology. *Proteins Struct. Funct. Genet.*, **23**, 415–421.
45. Olsson,M.H.M., Sondergaard,C.R., Rostkowski,M. and Jensen,J.H. (2011) PROPKA3: consistent treatment of internal and surface residues in empirical pK(a) predictions. *J. Chem. Theory Comput.*, **7**, 525–537.
46. Sondergaard,C.R., Olsson,M.H.M., Rostkowski,M. and Jensen,J.H. (2011) Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pK(a) values. *J. Chem. Theory Comput.*, **7**, 2284–2295.
47. Dolinsky,T.J., Nielsen,J.E., McCammon,J.A. and Baker,N.A. (2004) PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res.*, **32**, W665–W667.
48. Hornak,V., Abel,R., Okur,A., Strockbine,B., Roitberg,A. and Simmerling,C. (2006) Comparison of multiple amber force fields and development of improved protein backbone parameters. *Proteins: Struct. Funct. Bioinf.*, **65**, 712–725.
49. Meagher,K.L., Redman,L.T. and Carlson,H.A. (2003) Development of polyphosphate parameters for use with the AMBER force field. *J. Comput. Chem.*, **24**, 1016–1025.
50. Jorgensen,W.L., Chandrasekhar,J., Madura,J.D., Impey,R.W. and Klein,M.L. (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.*, **79**, 926–935.
51. Pronk,S., Pall,S., Schulz,R., Larsson,P., Bjelkmar,P., Apostolov,R., Shirts,M.R., Smith,J.C., Kasson,P.M., van der Spoel,D. *et al.* (2013) GROMACS 4.5: a high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics*, **29**, 845–854.
52. Van der Spoel,D., Lindahl,E., Hess,B., Groenhof,G., Mark,A.E. and Berendsen,H.J.C. (2005) Gromacs: fast, flexible, and free. *J. Comput. Chem.*, **26**, 1701–1718.
53. Essmann,U., Perera,L., Berkowitz,M.L., Darden,T., Lee,H. and Pedersen,L.G. (1995) A smooth particle mesh ewald method. *J. Chem. Phys.*, **103**, 8577–8593.
54. Brueckner,F. and Cramer,P. (2008) Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nat. Struct. Mol. Biol.*, **15**, 811–818.
55. Moran,S., Ren,R.X., Rumney,S. and Kool,E.T. (1997) Difluorotoluene, a nonpolar isostere for thymine, codes specifically and efficiently for adenine in DNA replication. *J. Am. Chem. Soc.*, **119**, 2056–2057.
56. Wang,L., Zhou,Y., Xu,L., Xiao,R., Lu,X., Chen,L., Chong,J., Li,H., He,C., Fu,X.-D. *et al.* (2015) Molecular basis for 5-carboxycytosine recognition by RNA polymerase II elongation complex. *Nature*, **523**, 621–625.
57. Westover,K.D., Bushnell,D.A. and Kornberg,R.D. (2004) Structural basis of transcription: separation of RNA from DNA by RNA polymerase II. *Science*, **303**, 1014–1016.
58. Kireeva,M., Kashlev,M. and Burton,Z.F. (2010) Translocation by multi-subunit RNA polymerases. *Biochim. Biophys. Acta*, **1799**, 389–401.
59. Maoileidigh,D.O., Tadigotla,V.R., Nudler,E. and Ruckenstein,A.E. (2011) A unified model of transcription elongation: what have we learned from single-molecule experiments? *Biophys J.*, **100**, 1157–1166.
60. Malinen,A.M., Turtola,M., Parthiban,M., Vainonen,L., Johnson,M.S. and Belogurov,G.A. (2012) Active site opening and closure control translocation of multisubunit RNA polymerase. *Nucleic Acids Res.*, **40**, 7442–7451.
61. Dangkulwanich,M., Ishibashi,T., Liu,S., Kireeva,M.L., Lubkowska,L., Kashlev,M. and Bustamante,C.J. (2013) Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. *Elife*, **2**, e00971.